



Harvard Business Review

REPRINT H04MX4
PUBLISHED ON HBR.ORG
NOVEMBER 05, 2018

ARTICLE **RESEARCH & DEVELOPMENT**

Using Experiments to Launch New Products

by Jeff Fossett, Duncan Gilchrist and Michael Luca

RESEARCH & DEVELOPMENT

Using Experiments to Launch New Products

by Jeff Fossett, Duncan Gilchrist and Michael Luca
NOVEMBER 05, 2018



MARTYN GODDARD/GETTY IMAGES

More than ever before, managers are using large-scale randomized controlled trials (i.e., experiments) to guide decisions. This has led to impressive gains for organizations ranging from Amazon to the UK government. We are excited about the rise of experiments in organizations and have spent much of the past few years thinking about how to design and interpret them. At the same time, we've seen that experimentation remains uneven across and within organizations, and many companies struggle with ways to start or expand experimenting.

One simple and often overlooked way for larger companies to experiment is to randomize the introduction of new products across a set of markets. To see how this can be valuable, consider how Uber rolled out its Express Pool service in 2018.

At the time, the company was already running UberPool, a service that allows passengers heading in the same direction to share rides and costs. With UberPool, passengers are picked up and dropped off wherever they like, as with other Uber services. But with the Express Pool service, which costs even less than UberPool, passengers are generally asked to walk short distances to meet their rides and to reach their destinations.

In 2018, in the run-up to the launch of Express Pool, Uber tasked one of us (Duncan, who manages a group of economists and data scientists within the company) with assessing how likely it was to succeed. How many riders would opt in, and how would the service affect the broader—and more complex—Uber ecosystem?

To answer those questions, Duncan and his team conducted an experiment, launching Express Pool in six large markets and then comparing metrics in the launch cities with those in others. Leveraging recent advances in experimental methods—especially a statistical method that allowed Uber to use a weighted combination of other cities to form a more-suitable “synthetic” control group—the team was able to tease out the ways in which the rollout was influencing Uber usage. Unsurprisingly, Express Pool created new kinds of trip matches. But the experiment also accounted for the effect that Express Pool had on existing Uber products and made clear that launching it would make good business sense. As a result, Uber was able confidently to introduce Express Pool to many of its major markets. This confidence, and the finding that inspired it, would not have been possible without the experiment.

Online marketplaces now abound, ranging from Uber and Airbnb to Rover and Tinder. And Uber is not alone among these companies in turning to market-level experiments to test new products and innovations.

Airbnb (where Jeff used to work as a data scientist) recently ran [an experiment](#) to test the impact of a new landing-page design on search-engine ranking and traffic. To run the experiment, Airbnb exploited the fact that it had landing pages with different URLs for different markets (San Francisco, Boston, New York, etc.). This meant that they could randomize the different URLs to include the new design or not, thereby isolating the design’s effect on search-engine traffic. And by doing that they were able to show that the new design was a success: the new landing page, it turned out, was driving a ~3.5% increase in search traffic, an improvement corresponding to tens of millions of incremental visitors per day for the platform. Based on these findings, Airbnb launched the new design for all markets.

It’s not just tech companies who can use large-scale experiments to test new products and innovations. Consider what a restaurant chain might do when deciding whether to offer a new

turkey-avocado sandwich. One traditional approach to a decision like this might be to roll out the sandwich in a couple of strategically chosen stores, run some focus groups, and study historical sales of other products. If people seem to like the sandwich in those stores, the business could roll it out in all of its stores and hope that it will succeed nationally. This type of approach would provide insight into the issue, but it has significant limitations. For example, it would be hard to know if the new sandwich crowded out other purchases. And it would be challenging to see whether this increased the overall number of customers. If the chain were to complement this approach with a large-scale randomized trial, by rolling the sandwich out to a set of randomly selected markets, they could learn much more about the effects that adding the sandwich might have broadly on sales (both for new and existing products), customer retention, and customer satisfaction.

We've seen companies miss important opportunities for experimentation, and we've seen experiments that suffer from implementation and interpretation challenges. For companies looking to test new products experimentally, here are some guidelines for getting started:

1) Decide what metrics matter to you most, and then come up with hypotheses about how they might behave. Invest in data collection and decide up-front what experimental outcomes will constitute success or failure. Create a map from data to decisions. Remember, it's great if more people buy a product, but not so great if this leads to more customer support calls.

2) Choose a random subset of markets (e.g., regions, cities, or franchises) in which to launch the product. The results of market-level experimentation are often noisy, so once you have a set of markets in mind, think carefully about whether you'll be able to detect the effects that you're hoping for. (Power calculations require a lot of assumptions but can help figure this out.)

3) Make sure to track not only whether your new product is working but also how its launch affects existing products. Express Pool on its own might look like a success, but if it doesn't sufficiently grow the overall market for rides, it's probably less valuable than it seems. Similarly, when Airbnb launches new products, the company needs to think about how bookings through existing products are affected. And when Starbucks rolled out its *sous vide* egg bites (if you haven't tried them, we recommend you do!), it needed to consider not only sales of the egg bites but also whether they crowded out other menu items.

4) Make sure you understand why your product is succeeding or failing. Top-line metrics like revenue and sales don't tell the whole story. Is the new product improving outcomes for some types of customers while harming them for others? Did the new product help one part of your acquisition funnel but hurt another? Do these moves align with your pre-experiment hypothesis? Understanding why a metric has moved can help you not only make a rollout decision but also understand how to innovate within a product space.

Jeff Fossett is a doctoral student at Harvard Business School, and previously worked as a data scientist at Airbnb

Duncan Gilchrist is a Data Science Manager at Uber Technologies. He holds a Ph.D. in Business Economics from Harvard University.

Michael Luca is the Lee J. Styslinger III Associate Professor of Business Administration at Harvard Business School.
